

Families in posets minimizing the number of comparable pairs

József Balogh*, Šárka Petríčková† and Adam Zsolt Wagner‡

March 17, 2017

Abstract

Given a poset P we say a family $\mathcal{F} \subseteq P$ is *centered* if it is obtained by ‘taking sets as close to the middle layer as possible’. A poset P is said to have the *centeredness property* if for any M , among all families of size M in P , centered families contain the minimum number of comparable pairs. Kleitman showed that the Boolean lattice $\{0, 1\}^n$ has the centeredness property. It was conjectured by Noel, Scott, and Sudakov, and by Balogh and Wagner, that the poset $\{0, 1, \dots, k\}^n$ also has the centeredness property, provided n is sufficiently large compared to k . We show that this conjecture is false for all $k \geq 2$ and investigate the range of M for which it holds. Further, we improve a result of Noel, Scott, and Sudakov by showing that the poset of subspaces of \mathbb{F}_q^n has the centeredness property. Several open questions are also given.

1 Introduction

Given a poset P , we say that two elements $A, B \in P$ form a *comparable pair* if $A \leq B$ or $B \leq A$. The study of families of sets containing few comparable pairs started with Sperner’s Theorem, a cornerstone result of combinatorics. It states that the largest antichain (i.e. family containing no comparable pairs) in the Boolean lattice $\mathcal{P}(n) = \{0, 1\}^n$ has size $\binom{n}{\lfloor n/2 \rfloor}$. The following natural question was first posed by Erdős and Katona for $r = 2$ and then extended by Kleitman [7] some fifty years ago: Given a poset $\mathcal{P}(n)$ and an integer M , what is the minimum number of r -chains that a family of M elements in $\mathcal{P}(n)$ must contain? For $r = 2$, the case of comparable pairs, the question was completely resolved by Kleitman [7]. For $r \geq 3$, we refer the reader to [3, 5, 6]. Here we are interested in the case $r = 2$, but for a general poset P .

Centered families in $\{0, 1, \dots, k\}^n$

We say that a family $\mathcal{F} \subseteq \{0, 1\}^n$ is *centered* if for any two sets $A, B \in \{0, 1\}^n$ with $A \in \mathcal{F}$ and $B \notin \mathcal{F}$ we have that

$$\left| |A| - \frac{n}{2} \right| \leq \left| |B| - \frac{n}{2} \right|,$$

where $|A|$ denotes the number of 1-coordinates in A . That is, \mathcal{F} is centered if it is constructed by “taking sets that are as close to the middle layer as possible”. This same notion can be extended

*Department of Mathematical Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA, jobal@math.uiuc.edu. The first author is partially supported by NSF Grant DMS-1500121, Arnold O. Beckman Research Award (UIUC Campus Research Board 15006) and by the Langan Scholar Fund (UIUC).

†University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA, petrckv2@illinois.edu.

‡University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA, zawagne2@illinois.edu.

to the poset $\{0, 1, \dots, k\}^n$ where $A \leq B$ if $A_i \leq B_i$ for all $i \in [n]$, where A_i and B_i are the i th coordinates of A and B . We say that a family $\mathcal{F} \subseteq \{0, 1, \dots, k\}^n$ is *centered* if for any two sets $A, B \in \{0, 1, \dots, k\}^n$ with $A \in \mathcal{F}$ and $B \notin \mathcal{F}$ we have that

$$\left| \sum_{i=1}^n A_i - \frac{nk}{2} \right| \leq \left| \sum_{i=1}^n B_i - \frac{nk}{2} \right|.$$

Denote by $\text{comp}(\mathcal{F})$ the number of comparable pairs in $\mathcal{F} \subseteq P$. A family $\mathcal{F} \subseteq P$ of size M is *M-optimal* if for all families $\mathcal{F}' \subseteq P$ of size M we have $\text{comp}(\mathcal{F}) \leq \text{comp}(\mathcal{F}')$. A poset $P \in \mathbf{P}$ has the *centeredness property* if for all $M \leq |P|$ there exists an M -optimal centered family. Using this terminology, Kleitman's celebrated theorem [7] can be stated as follows:

Theorem 1.1 (Kleitman [7], 1966). *The poset $\{0, 1\}^n$ has the centeredness property for all $n \in \mathbb{N}$.*

In [5] the authors characterised precisely which families achieve the minimum number of contained comparable pairs. It is natural to ask whether Theorem 1.1 holds for the poset $\{0, 1, \dots, k\}^n$ with $k \geq 2$ as well. It was showed in [3] that there exists a counterexample with $n = 2$ and $k = 16$. The following conjecture was raised independently in [9] and [3]:

Conjecture 1.2 (Noel–Scott–Sudakov [9], Balogh–Wagner [3]). *For every k there exists an n_0 such that if $n \geq n_0$ then the poset $\{0, 1, \dots, k\}^n$ has the centeredness property.*

Our main result is the construction of two different classes of explicit counterexamples to this natural generalisation of Theorem 1.1. We show that for every k , if n is sufficiently large, then there exists a suitable choice of M and a family \mathcal{F} of size M that contains strictly fewer comparable pairs than the centered families of the same size.

Denote by $\mathcal{L}_r(n, k)$ the r -th layer of $\{0, 1, \dots, k\}^n$, i.e. the set of vectors in $\{0, 1, \dots, k\}^n$ whose coordinates sum to r , and let $\ell_r(n, k) := |\mathcal{L}_r(n, k)|$. Write $\Sigma_r(n, k)$ for the total size of the r middle layers of $\{0, 1, \dots, k\}^n$. For $M \leq \Sigma_1(n, k)$ there exists an antichain of size M in the middle layer $\mathcal{L}_{\lfloor nk/2 \rfloor}(n, k)$ and hence Conjecture 1.2 trivially holds.

Our main result for the poset $\{0, 1, 2\}^n$ is the following.

Theorem 1.3. (a) *Let $\varepsilon > 0$, n be sufficiently large, and $M \leq (1 - \varepsilon)\Sigma_3(n, 2)$. Then there exists an M -optimal centered family in $\{0, 1, 2\}^n$.*

(b) *Let n be sufficiently large and $M = \Sigma_6(n, 2) - \binom{n}{3} - 1$. Then none of the centered families in $\{0, 1, 2\}^n$ are M -optimal.*

Theorem 1.3 says that the smallest $M = M_0$ for which Conjecture 1.2 breaks down (for $k = 2$) satisfies $(1 - \varepsilon)\Sigma_3(n, 2) < M_0 < \Sigma_6(n, 2) - \binom{n}{3}$. For $k = 2$ and M slightly larger than $\Sigma_1(n, 2)$ it was previously shown by Noel–Scott–Sudakov [9] that centered families contain asymptotically the optimal number of comparable pairs. They also obtained good lower bounds for the number of comparable pairs in larger families.

Theorem 1.4 (Noel–Scott–Sudakov [9]). *Let r be a fixed positive integer. Then there exists a constant $n_0(r)$ such that if $n \geq n_0(r)$ and $\mathcal{F} \subseteq \{0, 1, 2\}^n$ has cardinality at least $\Sigma_r(n, 2) + t$ then*

$$\text{comp}(\mathcal{F}) \geq \left(\frac{\ell_{3r-1}(n, 2)}{\ell_{2r-1}(n, 2)} - 1 \right) t.$$

While at first sight it may seem feasible that Conjecture 1.2 holds for much larger M , Theorem 1.5 (b) shows that this is not the case. Theorem 1.5 (a) provides a weak bound for the poset $\{0, 1, \dots, k\}^n$ that is optimal only for families not much larger than the middle layer.

Theorem 1.5. (a) For every $k \geq 2$, $n \geq 1$, and $t \geq 0$, every family $\mathcal{F} \subseteq \{0, 1, \dots, k\}^n$ of size $\Sigma_1(n, k) + t$ contains at least $t \lfloor \frac{n}{2} + 1 \rfloor$ comparable pairs.

(b) Let $k \geq 2$ and $\varepsilon > 0$. There exists a constant $n_0 = n_0(k, \varepsilon)$ such that for every $n \geq n_0$, if $M = \Sigma_j(n, k)$, where $(1 + \varepsilon) \log_2 n \leq j \leq \sqrt{n}/\log_2 n$, then none of the centered families in $\{0, 1, \dots, k\}^n$ are M -optimal.

Centered families in other posets

The notion of centeredness can be readily extended to several other common posets that satisfy some nice properties. In a poset P , y covers x if $x < y$ and there is no element z such that $x < z < y$. We say that the poset P is a *graded poset* if it is equipped with a rank function $\text{rk}: P \rightarrow \mathbb{N}$ which satisfies that $\text{rk}(x) < \text{rk}(y)$ whenever $x < y$, and $\text{rk}(y) = \text{rk}(x) + 1$ whenever y covers x . The *rank* of a poset P is the maximum rank of an element of P . Given a graded poset P , the r -th layer $\mathcal{L}_r(P)$ is the collection of elements in P of rank r , $\ell_r(P)$ is the size of $\mathcal{L}_r(P)$, and $\Sigma_r(P)$ is the total number of elements of P in the middle r layers. A graded poset of rank n is *rank-symmetric* if $\ell_i(P) = \ell_{n-i}(P)$ for $0 \leq i \leq n$ and it is *rank-unimodal* if $\ell_0(P) \leq \dots \leq \ell_j(P) \geq \ell_{j+1}(P) \geq \dots \geq \ell_n(P)$ for some $0 \leq j \leq n$. Denote by \mathbf{P} the family of all graded posets that are rank-symmetric and rank-unimodal, and by $\mathbf{P}(n)$ the posets in \mathbf{P} of rank n .

We will extend the notion of centeredness only to the posets in \mathbf{P} . Note that every $P \in \mathbf{P}(n)$ satisfies that its largest layer is $\mathcal{L}_{\lfloor n/2 \rfloor}(P)$ and its k largest layers are the k layers closest to the middle layer. Examples of such posets include $\{0, 1, \dots, k\}^n$ where $(A_1, \dots, A_n) \leq (B_1, \dots, B_n)$ if $A_i \leq B_i$ for all $1 \leq i \leq n$, and the poset $\mathcal{V}(q, n)$ of subspaces of \mathbb{F}_q^n ordered by inclusion where q is a prime power.

Similarly as before, given a poset $P \in \mathbf{P}(n)$, we say that a family $\mathcal{F} \subseteq P$ is *centered* if for any two sets $A, B \in P$ with $A \in \mathcal{F}$ and $B \notin \mathcal{F}$ we have that their ranks $\text{rk}(A), \text{rk}(B)$ satisfy

$$\left| \text{rk}(A) - \frac{n}{2} \right| \leq \left| \text{rk}(B) - \frac{n}{2} \right|.$$

In other words, \mathcal{F} is centered if it is constructed by “taking sets that are as close to the middle layer as possible”. Note that if $P = \{0, 1, \dots, k\}^n$, then this definition is the same as the definition of ‘centered’ introduced in the previous section (where the rank of P was nk).

Consider now for a prime power q the poset $\mathcal{V}(q, n)$ of subspaces of \mathbb{F}_q^n ordered by inclusion. Denote by $\begin{bmatrix} n \\ i \end{bmatrix}_q$ the number of subspaces of \mathbb{F}_q^n of dimension i . Note that $\begin{bmatrix} n \\ i \end{bmatrix}_q = \prod_{j=0}^{i-1} \frac{1-q^{n-j}}{1-q^{j+1}}$. The following result of Noel, Scott, and Sudakov [9] provides a lower bound on $\text{comp}(\mathcal{F})$ for $\mathcal{F} \subseteq \mathcal{V}(q, n)$.

Theorem 1.6 (Noel–Scott–Sudakov [9]). *Let q be a prime power and k be a fixed positive integer. There exists a constant $n_0(k)$ such that for $n \geq n_0(k)$ and $\mathcal{F} \subseteq \mathcal{V}(q, n)$,*

$$\text{If } |\mathcal{F}| \geq \sum_{r=0}^{k-1} \left[\begin{matrix} n \\ \lceil \frac{n-k+1+2r}{2} \rceil \end{matrix} \right]_q + t, \quad \text{then } \text{comp}(\mathcal{F}) \geq t \left[\begin{matrix} \lceil (n+k)/2 \rceil \\ k \end{matrix} \right]_q.$$

They pointed out that this bound is attained by a centered family and hence best possible when $k = 1$ and $0 \leq t \leq \left[\begin{matrix} n \\ \lfloor (n-1)/2 \rfloor \end{matrix} \right]_q$. We show that centered families are best for all sizes.

Theorem 1.7. *Let q be a prime power and $n \geq 1$. Then the poset $\mathcal{V}(q, n)$ has the centeredness property.*

Our proofs of Theorem 1.3 (a) and Theorem 1.7 are heavily based on the compression techniques of Kleitman [7]. The proof of Theorem 1.3 (b) arose when we attempted to prove that Conjecture 1.2 holds in the range $M \leq \Sigma_{(1-\varepsilon)\log_2 n}(n, 2)$ - all our proof attempts kept breaking down and they eventually led us to this counterexample. For Theorem 1.5 (a) we use the symmetric chain decomposition technique of Kleitman [8]. Finally the construction in Theorem 1.5 (b) came from the observation that for large enough M centered families are not even locally optimal, and in fact by replacing one of its elements in an appropriate way we can decrease the number of comparable pairs in the family.

For the corresponding maximization question, i.e. determining the maximum possible number of comparable pairs amongst families of size M in $\mathcal{P}(n)$ we refer the reader to [1].

2 Proof of Theorem 1.3 (a)

Whenever $A = (A_1, \dots, A_n)$ is an element of $\{0, 1, 2\}^n$, we will define the *size* (or *rank*) of A by $|A| := \sum_{i=1}^n A_i$. We will use a_0, a_1 and a_2 to denote the number of 0-, 1-, and 2-coordinates of A (that is, $a_i := |\{j : A_j = i\}|$). Similarly for $B \in \{0, 1, 2\}^n$ we will use the variables b_0, b_1, b_2 in the same fashion. The *complement* of a set $A \in \{0, 1, 2\}^n$ is defined as $A^c := (2 - A_1, \dots, 2 - A_n)$. For a permutation $\pi \in S_n$ and a set $A \in \{0, 1, 2\}^n$ we denote by $\pi(A)$ the set $(A_{\pi(1)}, \dots, A_{\pi(n)})$. For a family $\mathcal{F} \subseteq \{0, 1, 2\}^n$ and integer $0 \leq r \leq 2n$, we write $\mathcal{F}_r = \{A \in \mathcal{F} : |A| = r\}$ and $N_r(A) := \{B : |B| = r, B \subseteq A \text{ or } A \subseteq B\}$. Recall that in the poset $\{0, 1, 2\}^n$, $\mathcal{L}_r(n, 2)$ denotes the r -th layer and $\Sigma_j(n, 2)$ the total size of the j middle layers. In this section, we will often shorten $\mathcal{L}_r(n, 2)$ to \mathcal{L}_r and $\Sigma_j(n, 2)$ to Σ_j . Recall also that a family $\mathcal{F} \subseteq \{0, 1, 2\}^n$ of size M is called M -optimal if there is no other family $\mathcal{F}' \subseteq \{0, 1, 2\}^n$ of size M that contains strictly fewer comparable pairs than \mathcal{F} . Our goal is to show that there exists an M -optimal family that is centered.

Let $\varepsilon > 0$, let n be sufficiently large so that all the following estimates hold, and fix an $M \leq (1 - \varepsilon)\Sigma_3(n, 2)$. The proof is by induction on M , with the base case $M \leq \Sigma_1(n, 2)$ in which case there is an antichain in \mathcal{L}_n of size M and the claim follows. Hence we will assume that there exists an $(M - 1)$ -optimal centered family, and show that there exists an M -optimal centered family. Our first goal is to show that there exist M -optimal families that are contained in the middle three layers of $\{0, 1, 2\}^n$.

The following claim will be useful for us:

Claim 2.1. *Let $A, B \in \{0, 1, 2\}^n$ such that $B \subsetneq A$. If $|A|, |B| \geq n$, then for every $i \in \{1, \dots, |A| - |B|\}$:*

$$|N_{|B|+i}(B)| \leq |N_{|A|-i}(A)|.$$

Proof. Suppose that $|A|, |B| \geq n$. We show that B^c has at most as many 2's and at least as many 0's as A . This implies that there exists a permutation $\pi(B^c)$ of the coordinates of B^c such that $\pi(B^c) \subsetneq A$. Thus, $\pi(B^c)$ has at most as many neighbors in level $\mathcal{L}_{\pi(B^c)-i}$ as A does in level $\mathcal{L}_{|A|-i}$, for every $i \in \mathbb{N} \cup \{0\}$, so

$$|N_{|B|+i}(B)| = |N_{|B^c|-i}(B^c)| = |N_{|B^c|-i}(\pi(B^c))| \leq |N_{|A|-i}(A)|.$$

The number of 0's in B^c is equal to b_2 and the number of 2's in B^c is equal to b_0 . Hence we want to show that $b_0 \leq a_2$ and $b_2 \geq a_0$. Note first that since $B \subseteq A$, we have $b_2 \leq a_2$ and $b_0 \geq a_0$.

Let k, l be such that $|A| = n + k$ and $|B| = n + l$. From $|A| > |B| \geq n$ we have that $k > l \geq 0$. Since $a_0 + a_1 + a_2 = n$ and $a_1 + 2a_2 = n + k$, we have $a_2 - a_0 = k$, and similarly $b_2 - b_0 = l$.

$$b_0 = b_2 - l \leq a_2 - l \leq a_2 \quad \text{and} \quad b_2 = b_0 + l \geq b_0 \geq a_0.$$

□

A family \mathcal{F} in a poset $P \in \mathbf{P}$ is *compressed* if for every element $A \in \mathcal{F}$, every element comparable with A that is closer to the middle than A is in \mathcal{F} . Kleitman proved that every family in the Boolean lattice “can be compressed” without increasing the number of comparable pairs. It is not clear why this would be the case for $\{0, 1, \dots, k\}^n$ with $k > 2$. In the poset $\{0, 1, 2\}^n$ we can however at least obtain an analogous result for a weaker notion of top- and bottom-compressed, given in the following definition.

Definition 2.2. A family $\mathcal{F} \subseteq \{0, 1, 2\}^n$ is *top-compressed* if the following condition holds:

(T) If $A \in \mathcal{F}$ with $|A| > n$ and $B \subseteq A$ with $|B| \geq n$, then $B \in \mathcal{F}$.

A family $\mathcal{F} \subseteq \{0, 1, 2\}^n$ is *bottom-compressed* if the following condition holds:

(B) If $A \in \mathcal{F}$ with $|A| < n$ and $B \supseteq A$ with $|B| \leq n$, then $B \in \mathcal{F}$.

Lemma 2.3. *For every natural number $M \leq 3^n$, there exists an M -optimal family that is top- and bottom-compressed.*

Proof. Let \mathcal{F} be an M -optimal family. Suppose that there exist elements $A \in \mathcal{F}$ and $B \notin \mathcal{F}$ that violate condition (T). Pick such A for which $|A|$ is maximum, and then pick such B for which $|A| - |B|$ is minimal, and let $a = |A|$ and $b = |B|$. Then all elements in levels $\mathcal{L}_{b+1}, \dots, \mathcal{L}_{a-1}$ that are comparable with A are in \mathcal{F} .

We form a bipartite graph with parts $\mathcal{F}_a = \mathcal{F} \cap \mathcal{L}_a$ and $\overline{\mathcal{F}}_b = \mathcal{L}_b \setminus \mathcal{F}$ and with edges between comparable pairs. We write $N_{\mathcal{X}}(A)$ for the set of elements in \mathcal{X} comparable with A . Additionally, let $N_r(A) := N_{\mathcal{L}_r}(A)$, $N(A) := N_{\{0,1,2\}^n}(A)$, and $N_{\mathcal{X}}(\mathcal{A}) := \cup_{A \in \mathcal{A}} N_{\mathcal{X}}(A)$.

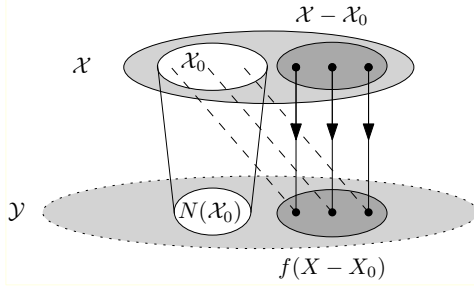
We will show that we can iteratively replace some elements of \mathcal{F}_a by elements of $\overline{\mathcal{F}}_b$ without increasing the number of comparable pairs. We will consider several cases based on sizes of \mathcal{F}_a and $\overline{\mathcal{F}}_b$ and the existence of “good” matchings that allow us to top-compress \mathcal{F} . Since $b < a$, the total value $\sum_{C \in \mathcal{F}} ||C| - n|$ of the family strictly decreases, ensuring that this process will terminate.

Suppose that we have families $\mathcal{A} \subseteq \mathcal{F}_a$ and $\mathcal{B} \subseteq \overline{\mathcal{F}}_b$ such that there is a perfect matching f between \mathcal{A} and \mathcal{B} . We define a new family $\mathcal{G} = (\mathcal{F} \setminus \mathcal{A}) \cup \mathcal{B}$ and show that the new family \mathcal{G} has no more comparable pairs than \mathcal{F} does. We compare the sizes of neighborhoods of \mathcal{A} and \mathcal{B} in the following four parts of the poset $\{0, 1, 2\}^n$:

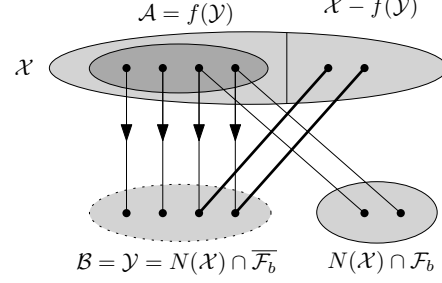
1. In levels $\mathcal{L}_{a+1}, \dots, \mathcal{L}_{2n}$: Since A is a greatest element of \mathcal{F} , no elements of \mathcal{F} are in these levels.
2. In levels $\mathcal{L}_0, \dots, \mathcal{L}_{b-1}$: Let $A \in \mathcal{A}$ and $B := f(A)$. Since $B \subseteq A$, if $C \subseteq B$ then $C \subseteq A$. So the number of comparable pairs cannot increase here.
3. In levels $\mathcal{L}_{b+1}, \dots, \mathcal{L}_{a-1}$: Since all elements in these levels are in \mathcal{F} , by Claim 2.1, for every $i \in [a - b - 1]$,

$$|\mathcal{F} \cap N_{b+i}(B)| \leq |N_{b+i}(B)| \leq |N_{a-i}(A)| = |\mathcal{F} \cap N_{a-i}(A)|.$$

Thus, every element $B \in \mathcal{B}$ has at most as many neighbors in $\mathcal{L}_{b+1} \cup \dots \cup \mathcal{L}_{a-1}$ as every $A \in \mathcal{A}$ does.



(a) Case 1.



(b) Case 2. If there exists a matching between \mathcal{X} and \mathcal{Y} covering \mathcal{Y} .

4. In levels \mathcal{L}_a and \mathcal{L}_b : This will be checked in each case separately.

In each case below, we present suitable sets $\mathcal{A} \in \mathcal{F}_a$ and $\mathcal{B} \in \overline{\mathcal{F}_b}$ with a perfect matching f between \mathcal{A} and \mathcal{B} for which

$$e(\mathcal{B}, \mathcal{G}_a) \leq e(\mathcal{A}, \mathcal{F}_b), \quad (2.1)$$

where $e(\mathcal{C}, \mathcal{D})$ denotes the number of edges between the families \mathcal{C} and \mathcal{D} .

Suppose first that there exists a matching f between \mathcal{F}_a and $\overline{\mathcal{F}_b}$ covering \mathcal{F}_a . Let $\mathcal{A} = \mathcal{F}_a$ and $\mathcal{B} = f(\mathcal{F}_a)$. Then there are no elements of \mathcal{G} in \mathcal{F}_a , so $e(\mathcal{B}, \mathcal{G}_a) = 0$. Henceforth we assume that there is no matching f between \mathcal{F}_a and $\overline{\mathcal{F}_b}$ covering \mathcal{F}_a , and we restrict our attention to the bipartite graph $(\mathcal{X}, \mathcal{Y})$, where

$$\mathcal{X} := \mathcal{F}_a \quad \text{and} \quad \mathcal{Y} := N(\mathcal{F}_a) \cap \overline{\mathcal{F}_b}.$$

Case 1: $|\mathcal{X}| \leq |\mathcal{Y}|$. By Hall's theorem, since there is no matching between \mathcal{X} and \mathcal{Y} covering \mathcal{X} , there must be a vertex set $\mathcal{X}_0 \subseteq \mathcal{X}$ such that $|N_{\mathcal{Y}}(\mathcal{X}_0)| < |\mathcal{X}_0|$. Choose \mathcal{X}_0 to be a maximal such vertex set. Then there must exist a matching f between $\mathcal{X} \setminus \mathcal{X}_0$ and $\mathcal{Y} \setminus N_{\mathcal{Y}}(\mathcal{X}_0)$ covering $\mathcal{X} \setminus \mathcal{X}_0$. Define $\mathcal{A} = \mathcal{X} \setminus \mathcal{X}_0$ and $\mathcal{B} = f(\mathcal{X} \setminus \mathcal{X}_0)$. Since there is no edge between $\mathcal{B} = f(\mathcal{X} \setminus \mathcal{X}_0)$ and $\mathcal{G}_a = \mathcal{X}_0$, the relation (2.1) holds.

Case 2: $|\mathcal{X}| > |\mathcal{Y}|$. Suppose first that there exists a matching f between \mathcal{X} and \mathcal{Y} covering \mathcal{Y} . Let $\mathcal{A} = f(\mathcal{Y})$ and $\mathcal{B} = \mathcal{Y}$. By Claim 2.1 applied with $i = a - b$ on every pair $(A, f(A)) \in (\mathcal{A}, \mathcal{B})$, we have $e(\mathcal{B}, \mathcal{L}_a) \leq e(\mathcal{A}, \mathcal{L}_b)$, so

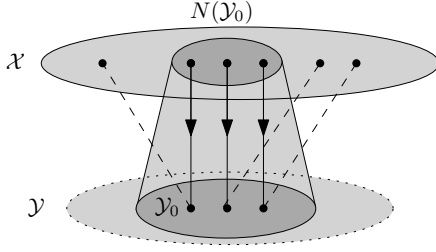
$$e(\mathcal{B}, \mathcal{G}_a) + e(\mathcal{B}, \mathcal{A}) = e(\mathcal{B}, \mathcal{F}_a) \leq e(\mathcal{B}, \mathcal{L}_a) \leq e(\mathcal{A}, \mathcal{L}_b) = e(\mathcal{A}, \mathcal{F}_b) + e(\mathcal{A}, \mathcal{B}).$$

The inequality (2.1) follows by subtracting $e(\mathcal{A}, \mathcal{B})$ on both sides. Suppose now that there is no matching covering \mathcal{Y} . By Hall's theorem, there must exist a minimal vertex set $\mathcal{Y}_0 \subseteq \mathcal{Y}$ such that $|N_{\mathcal{X}}(\mathcal{Y}_0)| < |\mathcal{Y}_0|$. Consider the following two subcases:

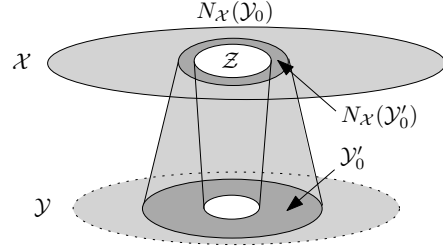
- There is a matching f between \mathcal{Y}_0 and $N_{\mathcal{X}}(\mathcal{Y}_0)$ covering $N_{\mathcal{X}}(\mathcal{Y}_0)$. Let $\mathcal{A} = N_{\mathcal{X}}(\mathcal{Y}_0)$ and $\mathcal{B} = f(N_{\mathcal{X}}(\mathcal{Y}_0))$. There is no edge between \mathcal{B} and $\mathcal{G}_a = \mathcal{F}_a \setminus \mathcal{A}$, hence $e(\mathcal{B}, \mathcal{G}_a) = 0$ and the inequality (2.1) trivially holds.
- There is no matching between \mathcal{Y}_0 and $N_{\mathcal{X}}(\mathcal{Y}_0)$ covering $N_{\mathcal{X}}(\mathcal{Y}_0)$. By Hall's theorem, there exists a vertex set $\mathcal{Z} \subseteq N_{\mathcal{X}}(\mathcal{Y}_0)$ with $|N_{\mathcal{Y}_0}(\mathcal{Z})| < |\mathcal{Z}|$. Then $\mathcal{Y}'_0 := \mathcal{Y}_0 \setminus N_{\mathcal{Y}_0}(\mathcal{Z})$ is smaller than \mathcal{Y}_0 . Since $|N_{\mathcal{X}}(\mathcal{Y}_0)| < |\mathcal{Y}_0|$ and $|\mathcal{Z}| > |N_{\mathcal{Y}_0}(\mathcal{Z})|$, we also have

$$|N_{\mathcal{X}}(\mathcal{Y}'_0)| \leq |N_{\mathcal{X}}(\mathcal{Y}_0)| - |\mathcal{Z}| < |\mathcal{Y}_0| - |N_{\mathcal{Y}_0}(\mathcal{Z})| = |\mathcal{Y}'_0|,$$

and we can conclude that \mathcal{Y}_0 was not a minimal set with $|N_{\mathcal{X}}(\mathcal{Y}_0)| < |\mathcal{Y}_0|$.



(c) Case 2a.



(d) Case 2b.

We showed that there exists an M -optimal family \mathcal{F} that is top-compressed. The proof that \mathcal{F} can “be made” bottom-compressed without increasing the number of comparable pairs follows by the above proof applied on $\mathcal{F}^c = \{A^c : A \in \mathcal{F}\}$. □

Lemma 2.3 ensures the existence of an M -optimal top- and bottom-compressed family. Although we will use the lemma only for $M \leq (1 - \varepsilon)\Sigma_3$, we emphasize that the result holds for any M , which might be of independent interest. Our next goal is to find an M -optimal family which additionally satisfies conditions (C1) and (C2) in the following definition.

Definition 2.4. We say that a family $\mathcal{F} \subseteq \{0, 1, 2\}^n$ of size M is *3-compressed* if \mathcal{F} is top-compressed, bottom-compressed, and additionally the following two conditions hold:

- (C1) If A is a maximal element of \mathcal{F} with $|A| = n + 2$ and $B \subseteq A$ is such that $|B| = n - 1$ and $b_0 > a_0$ then $B \in \mathcal{F}$.
- (C2) If A is a minimal element of \mathcal{F} with $|A| = n - 2$ and $B \supseteq A$ is such that $|B| = n + 1$ and $b_2 > a_2$ then $B \in \mathcal{F}$.

The following claim is an analogue statement to Claim 2.1.

Claim 2.5. Let $A, B \in \{0, 1, 2\}^n$ such that $B \subseteq A$. If $|A| = n + 2$, $|B| = n - 1$, and $b_0 \neq a_0$, then for every $i \in \{1, 2, 3\}$,

$$|N_{|B|+i}(B)| \leq |N_{|A|-i}(A)|.$$

Proof. Suppose that $|A| = n + 2$ and $|B| = n - 1$. Since $b_0 \neq a_0$, we only need to consider the following two cases:

Case 1: $b_2 = a_2$. The number of elements in levels $n + 1$, n , and $n - 1$, comparable with A , are

$$\alpha_1 := a_2 + a_1, \quad \alpha_2 := \binom{a_2 + a_1}{2} + a_2, \quad \text{and} \quad \alpha_3 := \binom{a_2 + a_1}{3} + a_2 \cdot (a_1 + a_2 - 1),$$

respectively. Similarly, the number of elements in levels n , $n + 1$, and $n + 2$, comparable with B , is

$$\beta_1 := b_0 + b_1, \quad \text{and} \quad \beta_2 := \binom{b_0 + b_1}{2} + b_0, \quad \text{and} \quad \beta_3 := \binom{b_0 + b_1}{3} + b_0 \cdot (b_1 + b_0 - 1),$$

respectively. Note that $a_1 = b_1 + 3$ and $a_2 = b_2 = b_0 - 1$, and so $a_2 + a_1 = b_0 + b_1 + 2$. We show that $\alpha_1 \geq \beta_1$, $\alpha_2 \geq \beta_2$, and $\alpha_3 \geq \beta_3$.

$$\begin{aligned}\alpha_1 - \beta_1 &= a_2 + a_1 - (b_0 + b_1) = b_0 + b_1 + 2 - (b_0 + b_1) > 0, \\ \alpha_2 - \beta_2 &= \binom{b_0+b_1+2}{2} + (b_0 - 1) - \left(\binom{b_0+b_1}{2} + b_0 \right) = 2(b_0 + b_1) \geq 0, \\ \alpha_3 - \beta_3 &= \binom{b_0+b_1+2}{3} + (b_0 - 1)(b_1 + b_0 + 1) - \left(\binom{b_0+b_1}{3} + b_0(b_1 + b_0 - 1) \right) \\ &= b_0^2 + 2b_0b_1 + b_1^2 + b_0 - b_1 - 1.\end{aligned}$$

The last expression is negative only if $b_0 = 0$ and $b_1 = 1$, which is not possible since every element $B \in \mathcal{L}_{n-1}$ must contain at least one 0-coordinate.

Case 2: $b_2 \leq a_2 - 1$ and $b_0 \geq a_0 + 1$. Then

$$b_0 = b_2 + 1 \leq a_2 \quad \text{and} \quad b_2 = b_0 - 1 \geq a_0.$$

So B^c has at most as many 2's and at least as many 0's as A , which implies that there exists a permutation $\pi(B^c)$ of the coordinates of B^c such that $\pi(B^c) \subseteq A$. This implies that for every $i \in \{1, 2, 3\}$,

$$|N_{|B|+i}(B)| = |N_{|B^c|-i}(B^c)| = |N_{|B^c|-i}(\pi(B^c))| \leq |N_{|A|-i}(A)|.$$

□

Lemma 2.6. *For every natural number $M \leq 3^n$, there exists an M -optimal family that is 3-compressed.*

Proof. Let \mathcal{F} be an M -optimal family in $\{0, 1, 2\}^n$ that is top- and bottom-compressed, whose existence is guaranteed by Lemma 2.3. If \mathcal{F} is not 3-compressed, then at least one of the conditions (C1) and (C2) fails. We assume that (C1) does not hold, keeping in mind that in the other case we can apply the same proof on \mathcal{F}^c . Suppose that there exists a comparable pair (A, B) in \mathcal{F} such that A is a maximal element with $|A| = n + 2$, $|B| = n - 1$, and $b_0 > a_0$. Let $a = |A|$ and $b = |B|$.

Let G be a bipartite graph with parts \mathcal{F}_a and $\overline{\mathcal{F}_b}$ and with edges between comparable pairs (A, B) for which $b_0 \neq a_0$. As in the proof of Lemma 2.3, we can iteratively replace some elements of \mathcal{F}_a by elements of $\overline{\mathcal{F}_b}$ without increasing the number of comparable pairs. We need to consider several cases based on sizes of \mathcal{F}_a and $\overline{\mathcal{F}_b}$ and existence of “good” matchings in G that allow us to compress \mathcal{F} . Since $b < a$, the total value $\sum_{C \in \mathcal{F}} ||C| - n|$ of the family strictly decreases, ensuring that this process will terminate. These cases are the same as in the proof of Lemma 2.3, except now we only consider matchings in the graph G (in which all pairs with $b_0 = a_0$ are removed), and we apply Claim 2.5 at every place we applied Claim 2.1 before.

□

We are almost ready to tackle Theorem 1.3 (a). We will need to make use of the fact that a typical set in $\{0, 1, 2\}^n$ of size n has about $n/3$ zeros $n/3$ ones, and $n/3$ twos.

Claim 2.7. *For every $\varepsilon > 10 \left(\frac{1}{1.1} \right)^{0.005n}$,*

$$\left| \left\{ A \in \mathcal{L}_{n+1} : \frac{0.9}{3}n \leq a_0 \leq \frac{1.1}{3}n \right\} \right| \geq (1 - \varepsilon^2)\ell_{n+1}.$$

Proof. For an integer $c \geq 0$, let $f(c) := |\{A \in \mathcal{L}_{n+1} : a_0 = c\}|$. Note that $f(c) = \binom{n}{c} \binom{n-c}{c+1}$, and hence

$$\frac{f(c)}{f(c+1)} = \frac{(c+1)(c+2)}{(n-2c-2)(n-2c-1)}.$$

If $c > \frac{1.07}{3}n$ we get $f(c)/f(c+1) > 1.1$ and if $c < \frac{0.93}{3}n = 0.31n$ we have $f(c)/f(c-1) > 1.1$. This means that

$$\sum_{i \leq 0.3n} f(i) \leq \frac{1}{1 - \frac{1}{1.1}} \cdot f(0.3n) \leq 11 \cdot \left(\frac{1}{1.1}\right)^{0.01n} \cdot f(0.31n) \leq \frac{\varepsilon^2}{2} \cdot \ell_{n+1}.$$

A similar computation gives $\sum_{i \geq 1.1n/3} f(i) \leq \frac{\varepsilon^2}{2} \cdot \ell_{n+1}$, and the claim follows. \square

The next claim shows that for slightly varying values of M , the M -optimal families contain about the same number of comparable pairs. For an integer N , write $\text{comp}(N)$ for the number of comparable pairs in an N -optimal family:

$$\text{comp}(N) := \min\{\text{comp}(\mathcal{F}) : \mathcal{F} \subseteq \{0, 1, 2\}^n, |\mathcal{F}| = N\}.$$

Claim 2.8. *If $M \leq (1 - \varepsilon) \sum_3(n, 2)$, then $\text{comp}(M) \leq \text{comp}(M - 1) + \frac{n^2}{4}$.*

Proof. By the induction hypothesis, there exists an $(M - 1)$ -optimal centered family \mathcal{G} . Since $M \leq (1 - \varepsilon) \sum_3$, the family \mathcal{G} consists of all elements in layer \mathcal{L}_n and some elements in layers \mathcal{L}_{n-1} and \mathcal{L}_{n+1} . Define

$$\mathcal{G}_1 := \{B \in \mathcal{L}_{n+1} : b_0 \geq \frac{0.9}{3}n\} \quad \text{and} \quad \mathcal{G}_2 := \{B \in \mathcal{L}_{n-1} : b_2 \geq \frac{0.9}{3}n\}.$$

Claim 2.7 implies $|\mathcal{G}_1|, |\mathcal{G}_2| \geq (1 - \varepsilon^2)|\mathcal{L}_{n+1}|$. For $M \leq (1 - \varepsilon) \sum_3$ we thus have $M < |\mathcal{L}_n| + |\mathcal{G}_1| + |\mathcal{G}_2|$. Add an element $B \in (\mathcal{G}_1 \cup \mathcal{G}_2) \setminus \mathcal{G}$ to \mathcal{G} . The element B is in at most $\binom{2.1n/3}{2} + n \leq \frac{n^2}{4}$ comparable pairs of $\mathcal{G} \cup \{B\}$, hence

$$\text{comp}(M) \leq \text{comp}(\mathcal{G} \cup \{B\}) \leq \text{comp}(\mathcal{G}) + \frac{n^2}{4} = \text{comp}(M - 1) + \frac{n^2}{4}.$$

\square

We are ready to finish the proof of Theorem 1.3 (a). Let \mathcal{F} be an M -optimal family that is 3-compressed, whose existence is guaranteed by Lemma 2.6, and assume that \mathcal{F} is not centered. This can mean one of two things:

1. The first possibility is that there exists an $A \notin \mathcal{F}$ of size $|A| = n$. Since \mathcal{F} is both top- and bottom-compressed, this means that there is no $B \in \mathcal{F}$ with $A \subseteq B$ or $B \subseteq A$, hence unless \mathcal{F} itself is an antichain we may decrease the number of comparable pairs in \mathcal{F} by replacing one of its elements by A .

2. The second possibility is that $\mathcal{L}_n \subseteq \mathcal{F}$ but $\mathcal{F} \not\subseteq \mathcal{L}_{n-1} \cup \mathcal{L}_n \cup \mathcal{L}_{n+1}$. Then there exists an element $A \in \mathcal{F}$ of size at least $n + 2$ or at most $n - 2$. By symmetry we may assume that there is an $A \in \mathcal{F}$ with $|A| \geq n + 2$. Since \mathcal{F} is 3-compressed, the number of elements in $\mathcal{F}_{a-1} \cup \mathcal{F}_{a-2} \cup \mathcal{F}_{a-3}$ comparable with A is at least

$$(a_1 + a_2) + \left(\binom{a_1 + a_2}{2} + a_2\right) + \left(\binom{a_1 + a_2}{3} + a_2(a_1 + a_2 - 1) - \binom{a_2}{3}\right). \quad (2.2)$$

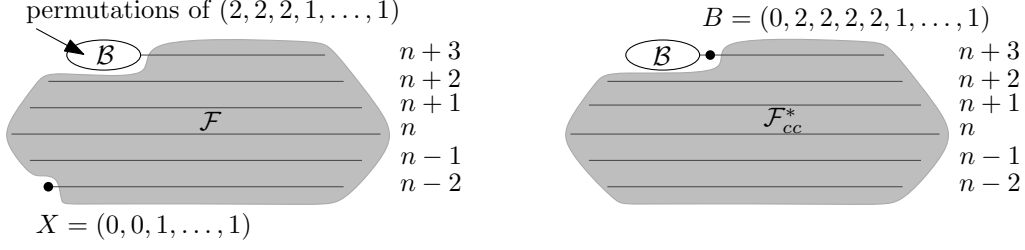


Figure 1: A non-centered family \mathcal{F} and a canonical centered family \mathcal{F}_{cc}^* such that $\text{comp}(\mathcal{F}) < \text{comp}(\mathcal{F}_{cc}^*) \leq \text{comp}(\mathcal{F}_{cc})$ for every canonical centered family \mathcal{F}_{cc} of size $M = \Sigma_6 - \binom{n}{3} - 1$.

The term $\binom{a_2}{3}$ accounts for the elements of \mathcal{L}_{n-1} comparable with A that have a_0 zeros, which are not necessarily in \mathcal{F} by the definition of 3-compressed. Observe that every such element is formed by decreasing three 2-coordinates of A to 1-coordinates, giving $\binom{a_2}{3}$ choices. Since $a_1 + a_2 \geq n + 2 - a_2$, the quantity (2.2) is minimized when $a_1 = 0$ and $a_2 = \frac{n+2}{2}$. It follows that this quantity is at least $a_2 + \binom{a_2}{2} + a_2^2 \geq \frac{3}{2}a_2^2 \geq \frac{3}{8}n^2$. But then $\text{comp}(\mathcal{F}) > \text{comp}(\mathcal{F} \setminus \{A\}) + \frac{3n^2}{8}$, and \mathcal{F} was not M -optimal (by Claim 2.8), a contradiction. \square

3 Proof of Theorem 1.3 (b)

Recall that for an integer $a \geq 0$ and a family $\mathcal{G} \subseteq \{0, 1, 2\}^n$ we have the notation $\mathcal{G}_a = \{A \in \mathcal{G} : |A| = a\}$ and $\mathcal{L}_a = \mathcal{L}_a(n, 2)$. We say that a centered family $\mathcal{G} \subseteq P$ is *canonical centered* if there exists at most one $\ell \geq 0$ with $0 < |\mathcal{G}_\ell| < |\mathcal{L}_\ell(P)|$, i.e. if it has at most one partial layer (while centered families could have two). As in Section 2, whenever A and B are elements of $\{0, 1, 2\}^n$, we write a_0, a_1, a_2 and b_0, b_1, b_2 for the number of 0-, 1-, 2-coordinates in A and B respectively. For an element $A \in \{0, 1, 2\}^n$ and family $\mathcal{G} \subseteq \{0, 1, 2\}^n$, we use the notation $\text{comp}(A, \mathcal{G}) := |\{B \in \mathcal{G} : B \subsetneq A \text{ or } A \subsetneq B\}|$ and $\text{Comp}(\mathcal{G}) := \{(A, B) \in \mathcal{G} \times \mathcal{G} : A \subset B\}$, so that $|\text{Comp}(\mathcal{G})| = \text{comp}(\mathcal{G})$.

Let $X = (0, 0, 1, 1, \dots, 1) \in \mathcal{L}_{n-2}$, $\mathcal{B} = \{B \in \mathcal{L}_{n+3} : b_0 = 0\}$, and $\mathcal{C} := \{C \in \{0, 1, 2\}^n : n - 2 \leq |C| \leq n + 3\}$. Finally, let $\mathcal{F} := \mathcal{C} \setminus (\mathcal{B} \cup \{X\})$ (see Figure 1). Then \mathcal{F} is not centered, but we claim that \mathcal{F} contains fewer comparable pairs than every centered family of size $M = |\mathcal{F}| = \Sigma_6(n, 2) - \binom{n}{3} - 1$. The proof of this claim goes in two stages. First we show that \mathcal{F} contains fewer comparable pairs than the best canonical centered family of this size (Claim 3.1), and next we show that among centered families of this size the canonical families are the best (Lemma 3.2).

Claim 3.1. *Whenever \mathcal{F}_{cc} is a canonical centered family of size $M = \Sigma_6(n, 2) - \binom{n}{3} - 1$ we have $\text{comp}(\mathcal{F}) < \text{comp}(\mathcal{F}_{cc})$.*

Proof. Every canonical centered family \mathcal{F}_{cc} of size $M = \Sigma_6(n, 2) - \binom{n}{3} - 1$ consists of all elements in levels $\mathcal{L}_{n-2}, \dots, \mathcal{L}_{n+2}$ and $\ell_{n+3} - \binom{n}{3} - 1$ elements in \mathcal{L}_{n+3} (or $\ell_{n-3} - \binom{n}{3} - 1$ elements in \mathcal{L}_{n-3} , in which case the proof is symmetrical). Let $B = (0, 2, 2, 2, 2, 1, 1, \dots, 1) \in \mathcal{L}_{n+3}$ and note that $\mathcal{F}_{cc}^* := \mathcal{F} \cup \{X\} \setminus \{B\}$ is one of the canonical centered families of size M with the least number of contained comparable pairs. Indeed, removing all elements with no 0-coordinates plus one element with one 0-coordinate from \mathcal{L}_{n+3} ensures the smallest possible number of comparable pairs. This

can be seen because it is always better to replace a 2-coordinate and a 0-coordinate by two 1-coordinates, or directly from the formula (2.2).

It suffices to show that $\text{comp}(B, \mathcal{F}) < \text{comp}(X, \mathcal{F})$ since then we can improve $\mathcal{F} \cup \{X\} \setminus \{B\}$ by deleting X and adding B . Now, $\text{comp}(X, \mathcal{F}) \geq \binom{n}{5} + \binom{n}{4}$ whereas $\text{comp}(B, \mathcal{F}) = \binom{n-1}{5} + \binom{n-1}{4} + O(n^3)$, which is $\Theta(n^4)$ smaller than $\text{comp}(X, \mathcal{F})$ and the claim follows. \square

Lemma 3.2. *Among centered families of size $M = \Sigma_6(n, 2) - \binom{n}{3} - 1$ the function $\text{comp}(\cdot)$ attains its minimum on a canonical centered family.*

Proof. Define a partial order on the collection of centered families of size M by letting $\mathcal{H} < \mathcal{H}'$ if $\text{comp}(\mathcal{H}) < \text{comp}(\mathcal{H}')$, or if $\text{comp}(\mathcal{H}) = \text{comp}(\mathcal{H}')$ and $|\mathcal{H}_{n+3}| > |\mathcal{H}'_{n+3}|$. We will show that one of the minimal elements of this partial order is canonical centered, which immediately implies Lemma 3.2. Let \mathcal{G} be a centered family of size $M = \Sigma_6(n, 2) - \binom{n}{3} - 1$ that is minimal according to this ordering. Note that $\mathcal{L}_{n-2} \cup \dots \cup \mathcal{L}_{n+2} \subseteq \mathcal{G} \subseteq \mathcal{L}_{n-3} \cup \dots \cup \mathcal{L}_{n+3}$.

Given a permutation $\pi \in S_n$ of order 2 (i.e., $\pi^2 = 1$) define the π -compression of \mathcal{G} by “replace $A \in \mathcal{G}_{n-3}$ by $\pi(A^c)$ unless it is already in \mathcal{G}_{n+3} ”. That is,

$$\text{cpr}_\pi(\mathcal{G}) = \mathcal{G} \cup \{\pi(A^c) \in \mathcal{L}_{n+3} : A \in \mathcal{G}_{n-3}\} \setminus \{A \in \mathcal{G}_{n-3} : \pi(A^c) \notin \mathcal{G}_{n+3}\}.$$

Claim 3.3. *For every $\pi \in S_n$ of order 2 we have $\text{cpr}_\pi(\mathcal{G}) < \mathcal{G}$, unless $\mathcal{G} = \text{cpr}_\pi(\mathcal{G})$. That is, π -compression improves the family unless it is already π -compressed.*

Proof. Note first that unless $\mathcal{G} = \text{cpr}_\pi(\mathcal{G})$ we have that $|\text{cpr}_\pi(\mathcal{G})_{n+3}| > |\mathcal{G}_{n+3}|$. It thus remains to show $\text{comp}(\text{cpr}_\pi(\mathcal{G})) \leq \text{comp}(\mathcal{G})$. Suppose that $B \subset \pi(A^c)$ is a new comparable pair. Then A was replaced by $\pi(A^c)$, so $A \in \mathcal{G} \setminus \text{cpr}_\pi(\mathcal{G})$. The element B was not replaced by $\pi(B^c)$, so $\pi(B^c) \in \mathcal{G}$. Observe that for every $\pi \in S_n$, $B \subset \pi(A^c)$ implies $A \subset \pi^{-1}(B^c)$. Since our permutation π is of order 2, we have $\pi^{-1}(B^c) = \pi(B^c)$, and thus $A \subset \pi(B^c)$. Together, for every new comparable pair $B \subset \pi(A^c)$ there is an old comparable pair $A \subset \pi(B^c)$ which got deleted during the compression. This defines an injection from $\text{Comp}(\text{cpr}_\pi(\mathcal{G})) \setminus \text{Comp}(\mathcal{G})$ into $\text{Comp}(\mathcal{G}) \setminus \text{Comp}(\text{cpr}_\pi(\mathcal{G}))$ and the claim follows. \square

We sketch the idea of the remaining part of the proof. By Claim 3.3 and the minimality of \mathcal{G} , the family \mathcal{G} is π -compressed for all permutations π of order 2. For $A \in \mathcal{G}_{n-3}$, define

$$\Pi(A^c) := \{\pi(A^c) \in \mathcal{L}_{n+3} : \pi \in S_n \wedge \pi^2 = 1\},$$

and count the elements of $\Pi(A^c)$ comparable with A . Every such element has to be in \mathcal{G}_{n+3} by definition of π -compression. To obtain a superset of A in $\Pi(A^c)$, we first need to switch all 0-coordinates of A^c with some of its 2-coordinates. After that we can freely switch any of the remaining three 2-coordinates with any three 1-coordinates. Any permutation that is formed in this fashion is obviously of order 2. The number of such permutations is $\binom{a_0}{3} \binom{a_1}{3}$. It follows that if the number of 0's and 1's in A is (close to) linear in n , then the number of elements in \mathcal{G}_{n+3} comparable with A is of order (close to) n^6 . Therefore, \mathcal{G}_{n-3} cannot have many such elements since otherwise we could replace \mathcal{G}_{n-3} by elements of $\mathcal{L}_{n+3} \setminus \mathcal{G}$ and the number of comparable pairs would decrease. We partition \mathcal{G} into \mathcal{G}' , \mathcal{G}'' , and \mathcal{G}^* as follows:

$$\mathcal{G}' = \left\{ A \in \mathcal{G}_{n-3} : a_2 \leq n^{2/3} \log n \right\}, \quad \mathcal{G}'' = \left\{ A \in \mathcal{G}_{n-3} : a_2 \geq \frac{n}{2} - n^{2/3} \log n \right\}, \quad \mathcal{G}^* = \mathcal{G}_{n-3} - (\mathcal{G}' \cup \mathcal{G}'').$$

Observe that \mathcal{G}' contains elements with a small number of 0- and 2-coordinates while \mathcal{G}'' contains elements with small number of 1-coordinates. Claim 3.4 states that there cannot be more elements in \mathcal{G}^* than in $\mathcal{G}' \cup \mathcal{G}''$. Claim 3.5 uses a similar averaging argument to bound $|\mathcal{G}' \cup \mathcal{G}''|$ by $2|\mathcal{H}' \cup \mathcal{H}''|$, where $\mathcal{H}' \cup \mathcal{H}''$ is the family of sets in $\mathcal{G}' \cup \mathcal{G}''$ that are in a small number of comparable pairs in \mathcal{G} . Claim 3.6 then implies that $\mathcal{H}' \cup \mathcal{H}''$ must be empty, and we conclude that \mathcal{G} is canonical centered.

Claim 3.4. $|\mathcal{G}_{n-3}| \leq 2|\mathcal{G}' \cup \mathcal{G}''|$.

Proof. Let A be an element of \mathcal{G}^* and consider all its supersets of the form $\pi(A^c)$ with $\pi^2 = 1$. Since \mathcal{G} is π -compressed for every involution π , we know that all these supersets are in \mathcal{G} . Let Π_A be the set of a permutations π of order 2 such that each π switches all 0-coordinates of A^c with all but three of its 2-coordinates, and the remaining three 2-coordinates with three arbitrary 1-coordinates. Equivalently, for every $\pi \in \Pi_A$, the element $\pi(A^c)$ is formed from A by increasing three 0-coordinates and three 1-coordinates by one. We thus always have $A \subset \pi(A^c)$, and hence the number of supersets of A in $\Pi(A^c)$ is at least $|\Pi_A| = \binom{a_0}{3} \binom{a_1}{3}$. Since $A \notin \mathcal{G}' \cup \mathcal{G}''$ and $a_0 = a_2 + 3$, we have

$$n^{2/3} \log n + 3 \leq a_0 \leq \frac{n}{2} - n^{2/3} \log n + 3.$$

From $a_0 + a_1 + a_2 = n$ we have $a_1 = n - 2a_0 + 3$, and thus

$$a_1 \geq n - 2 \left(\frac{n}{2} - n^{2/3} \log n + 3 \right) + 3 = 2n^{2/3} \log n - 3.$$

As either a_0 or a_1 is larger than $n/10$, we have

$$|\Pi_A| = \binom{a_0}{3} \binom{a_1}{3} \geq n^5 \log^2 n.$$

We claim that the elements of \mathcal{G}_{n-3} are in at most n^5 comparable pairs each on average. Indeed, otherwise we could replace \mathcal{G}_{n-3} by an arbitrary subset of $\overline{\mathcal{G}_{n+3}} = \mathcal{L}_{n+3} \setminus \mathcal{G}_{n+3}$ of size $|\mathcal{G}_{n-3}|$ and obtain a canonical centered family with a smaller number of comparable pairs. Because each element of \mathcal{G}^* is in at least $n^5 \log^2 n$ comparable pairs, we have $|\mathcal{G}^*| \leq |\mathcal{G}' \cup \mathcal{G}''|$, and the claim follows. \square

Let

$$\mathcal{H}' = \{A \in \mathcal{G}' : \text{comp}(A, \mathcal{G}) \leq 2n^5\} \quad \text{and} \quad \mathcal{H}'' = \{A \in \mathcal{G}'' : \text{comp}(A, \mathcal{G}) \leq 2n^5\}.$$

Claim 3.5. $|\mathcal{G}' \cup \mathcal{G}''| \leq 2|\mathcal{H}' \cup \mathcal{H}''|$.

Proof. As before, the elements of \mathcal{G}_{n-3} must be in at most n^5 comparable pairs each on average since otherwise we could replace \mathcal{G}_{n-3} by an arbitrary subset of $\overline{\mathcal{G}_{n+3}}$. Recall that the family \mathcal{G}_{n-3} is partitioned into \mathcal{G}' , \mathcal{G}'' and \mathcal{G}^* , and that every element of \mathcal{G}^* is in at least $n^5 \log^2 n$ comparable pairs of \mathcal{G} (see proof of Claim 3.4). We thus necessarily have $|\mathcal{G}' \cup \mathcal{G}''| \leq 2|\mathcal{H}' \cup \mathcal{H}''|$. \square

Claim 3.6.

$$|\mathcal{H}'|, |\mathcal{H}''| \leq \frac{\log^8 n}{n^2} \cdot |\overline{\mathcal{G}_{n+3}}|.$$

Proof. We first count the number E'' of comparable pairs $(A, B) \in \mathcal{H}'' \times \overline{\mathcal{G}_{n+3}}$ such that $a_2 = b_2$. We count E'' two ways:

1. Let $A \in \mathcal{H}'' \subseteq \mathcal{G}''$. Then $a_0 = a_2 + 3 \geq \frac{n}{2} - n^{2/3} \log n + 3$ by the definition of \mathcal{G}'' . We need to count the number of sets $B \in \overline{\mathcal{G}_{n+3}}$ formed from A by increasing six of its 0-coordinates to 1-coordinates. Since $\text{comp}(A, \mathcal{G}_{n+3}) \leq 2n^5$ by the definition of \mathcal{H}'' , this number is at least $\binom{a_0}{6} - 2n^5 \geq \binom{n/3}{6} \geq n^6/10^9$.
2. Let now $B \in \overline{\mathcal{G}_{n+3}}$ for which there exists an $A \in \mathcal{H}''$ with $a_2 = b_2$. Then

$$b_1 = n + 3 - 2b_2 = n + 3 - 2a_2 \leq n + 3 - 2\left(\frac{n}{2} - n^{2/3} \log n\right) \leq 3n^{2/3} \log n.$$

Therefore, the number of sets A formed from B by decreasing six of its 1-coordinates to 0-coordinates is at most $\binom{3n^{2/3} \log n}{6} \leq n^4 \log^7 n$.

Together we obtain

$$|\mathcal{H}''| \cdot \frac{n^6}{10^9} \leq E'' \leq |\overline{\mathcal{G}_{n+3}}| \cdot n^4 \log^7 n, \quad (3.1)$$

and the second inequality in Claim 3.6 follows.

Similarly, we count the number E' of comparable pairs $(A, B) \in \mathcal{H}' \times \overline{\mathcal{G}_{n+3}}$ such that $a_0 = b_0$.

1. Let $A \in \mathcal{H}' \subseteq \mathcal{G}'$. Then $a_1 = n - 3 - 2a_2 \geq n - 3 - 2n^{2/3} \log n$ by the definition of \mathcal{G}' . The number of sets $B \in \overline{\mathcal{G}_{n+3}}$ formed from A by increasing six of its 1-coordinates to 2-coordinates is at least $\binom{a_1}{6} - 2n^5 \geq \binom{n/3}{6} \geq n^6/10^9$.
2. Let now $B \in \overline{\mathcal{G}_{n+3}}$ for which there exists an $A \in \mathcal{H}'$ with $a_0 = b_0$. Then $b_2 = a_2 + 6 \leq n^{2/3} \log n + 6$. Therefore, the number of sets A formed from B by decreasing six of its 1-coordinates to 0-coordinates is at most $\binom{n^{2/3} \log n + 6}{6} \leq n^4 \log^7 n$.

Similarly to 3.1 we have

$$|\mathcal{H}'| \cdot \frac{n^6}{10^9} \leq E' \leq |\overline{\mathcal{G}_{n+3}}| \cdot n^4 \log^7 n,$$

and the first inequality in Claim 3.6 follows. \square

We are ready to finish the proof of Lemma 3.2. Applying the previous three claims, we obtain

$$|\mathcal{G}_{n-3}| \stackrel{C3.4}{\leq} 2|\mathcal{G}' \cup \mathcal{G}''| \stackrel{C3.5}{\leq} 4|\mathcal{H}' \cup \mathcal{H}''| \stackrel{C3.6}{\leq} \frac{\log^9 n}{n^2} \cdot |\overline{\mathcal{G}_{n+3}}| = \frac{\log^9 n}{n^2} \left(|\mathcal{G}_{n-3}| + \binom{n}{3} + 1 \right),$$

and therefore

$$|\mathcal{G}_{n-3}| \leq n \log^{10} n. \quad (3.2)$$

Assume that $\mathcal{H}'' \neq \emptyset$ and let $A \in \mathcal{H}''$. As in the proof of Claim 3.6, $\text{comp}(A, \overline{\mathcal{G}_{n+3}}) \geq \binom{a_0}{6} - 2n^5 \geq n^6/10^9$, and so $|\overline{\mathcal{G}_{n+3}}| \geq n^6/10^9$. This implies $|\mathcal{G}_{n-3}| \geq n^6/10^{10}$, which contradicts equation (3.2). By the same argument we have $\mathcal{H}' = \emptyset$. Hence $\mathcal{G}_{n-3} = \emptyset$ by Claims 3.4 and 3.5, and we conclude that \mathcal{G} is canonical centered, proving the lemma. \square

4 Proof of Theorem 1.5 (a)

A *symmetric chain* \mathcal{A} in the poset $\{0, 1, \dots, k\}^n$ is a chain $A^1 \subset A^2 \subset \dots \subset A^\ell$ such that $|A^{i+1}| = |A^i| + 1$ for all $i < \ell$, and $|A^1| + |A^\ell| = nk$. A *symmetric chain decomposition* (or SCD in short) of the poset $\{0, 1, \dots, k\}^n$ is a decomposition of the poset into disjoint symmetric chains. It is known that there exists an SCD of $\{0, 1, \dots, k\}^n$, see e.g. [2, Theorem 3.6.1].

Define

$$\delta_{n,k}(B, a) := |\{A \in \{0, 1, \dots, k\}^n : A \subset B, |A| = a\}|,$$

and let, for $0 < a < b \leq nk$

$$\delta_{n,k}(b, a) := \min_{|B|=b} \delta_{n,k}(B, a).$$

For a set $A = (A_1, \dots, A_n) \in \{0, 1, \dots, k\}^n$, define its *complement* by $A^c := (k - A_1, \dots, k - A_n)$.

Following the ideas of [8], given a comparable pair $A \subset B$ with $A, B \in \{0, 1, \dots, k\}^n$, define their *weight* $\omega(A, B)$ to be equal to the probability that A, B are contained in the same chain of the SCD \mathcal{X} , if \mathcal{X} is chosen uniformly at random from the collection of all SCDs of $\{0, 1, \dots, k\}^n$. Notice that if $A \subset B$ is a comparable pair with $||A| - \frac{nk}{2}| \leq ||B| - \frac{nk}{2}|$ then

$$\omega(A, B) = \delta_{n,k}(B, |A|)^{-1} \leq \delta_{n,k}(|B|, |A|)^{-1}$$

and if on the other hand we have $||A| - \frac{nk}{2}| \geq ||B| - \frac{nk}{2}|$ then

$$\omega(A, B) = \delta_{n,k}(A^c, nk - |B|)^{-1} \leq \delta_{n,k}(nk - |A|, nk - |B|)^{-1}.$$

Claim 4.1. *For any comparable pair $A \subset B$ we have $\omega(A, B) \leq \lfloor \frac{n}{2} + 1 \rfloor^{-1}$, unless $A = \emptyset$ and $B = (k, k, \dots, k)$.*

Proof. If $A = \emptyset$ and $B = (k, k, \dots, k)$, then $\omega(A, B) = 1$, so assume that this is not the case. W.l.o.g. we may further assume that $||A| - \frac{nk}{2}| \leq ||B| - \frac{nk}{2}|$, so that $\omega(A, B) \leq \delta_{n,k}(|B|, |A|)^{-1}$, and hence it suffices to prove that

$$\delta_{n,k}(|B|, |A|) \geq \left\lfloor \frac{n}{2} + 1 \right\rfloor \quad (4.1)$$

holds whenever $||A| - \frac{nk}{2}| \leq ||B| - \frac{nk}{2}|$. From now on we will assume that nk is even – the proof for the odd case is very similar. Note that for fixed n, k, a , the function $\delta_{n,k}(b, a)$ is an increasing function of b , hence it suffices to show (4.1) in the following cases:

- (1) $|B| = |A| + 1$, and
- (2) $\frac{kn}{2} - |A| = |B| - \frac{kn}{2}$.

The first case is easier: if $|B| = |A| + 1$ and $||A| - \frac{nk}{2}| \leq ||B| - \frac{nk}{2}|$ then B has at least $\lfloor \frac{n}{2} + 1 \rfloor$ non-zero coordinates and hence $\delta_{n,k}(B, |A|) \geq \lfloor \frac{n}{2} + 1 \rfloor$ as required.

Assume now that we are in the second case and not in the first case, and so $\frac{kn}{2} - |A| = |B| - \frac{kn}{2} \geq 1$. We construct an appropriate B^* with $B^* \subseteq B$, using the following algorithm:

1. Set $B^* := B$.
2. Pick an $i \in [n]$ such that $B_i^* \notin \{0, 1\}$. Decrease B_i^* by one.
3. If all coordinates of B^* are zero or one then STOP.
4. If $|B^*| = |A| + 1$ then STOP.
5. Go to Step 2.

Notice that if we terminated in Step 3 then we have

$$\delta_{n,k}(B, |A|) \geq \delta_{n,k}(B^*, |A|) = \binom{|B^*|}{|A|} \geq |B^*| \geq \left\lfloor \frac{n}{2} + 1 \right\rfloor,$$

where the second-to-last inequality follows since $A \neq \emptyset$ and the last inequality follows from the fact that B^* has the same number of non-zero coordinates as B , which is at least $\lfloor \frac{n}{2} + 1 \rfloor$ as $|B| > nk/2$.

On the other hand, if we terminated in Step 4 then we have, similarly as above, that

$$\delta_{n,k}(B, |A|) \geq \delta_{n,k}(B^*, |A|) = |B^*| \geq \left\lfloor \frac{n}{2} + 1 \right\rfloor,$$

and the proof is complete. \square

Using this claim we are now ready to prove Theorem 1.5 (a). Say that an SCD $\mathcal{X} = \{X_1, \dots, X_m\}$ contains a comparable pair A, B if there exists an $i \in [m]$ such that $A, B \in X_i$. Given a family \mathcal{F} , write $N(\mathcal{F}, \mathcal{X})$ for the number of comparable pairs in \mathcal{F} that are contained in \mathcal{X} .

Proof of Theorem 1.5 (a). If $M = (k+1)^n$ then $\mathcal{F} = \{0, 1, \dots, k\}^n$ and the theorem holds, otherwise we may assume that $\emptyset \notin \mathcal{F}$. Since the largest antichain in $\{0, 1, \dots, k\}^n$ has size $\Sigma_1(n, k)$, by the pigeonhole principle every SCD contains at least x comparable pairs. The probability that a randomly chosen SCD contains a fixed comparable pair A, B is by the definition $\omega(A, B)$. Hence the expected number of comparable pairs contained by a random SCD satisfies

$$x \leq \mathbf{E}_{\mathcal{X}}(N(\mathcal{F}, \mathcal{X})) = \sum_{A, B \in \mathcal{F}, A \subseteq B} \omega(A, B) \leq \text{comp}(\mathcal{F}) \left(\left\lfloor \frac{n}{2} + 1 \right\rfloor \right)^{-1},$$

where the last inequality follows from Claim 4.1, and the proof is complete. \square

5 Proof of Theorem 1.5 (b)

Let $P = \{0, 1, \dots, k\}^n$ where k is a fixed constant, $0 < \varepsilon < 0.01$, and n be sufficiently large so that all following estimates hold. We are given an integer j with $(1 + \varepsilon) \log_2 n \leq j \leq \sqrt{n}/\log_2 n$ and we have $M = \Sigma_j(n, k)$. For simplicity we will assume $nk + j$ is even, the odd case is very similar, and we omit the details. Let

$$\mathcal{F} := \left\{ A \in P : \frac{nk - j}{2} < |A| \leq \frac{nk + j}{2} \right\}.$$

Let B be such that $|B| = \frac{nk+j}{2}$ and every coordinate of B is either $\lfloor \frac{k}{2} \rfloor$ or $\lfloor \frac{k}{2} \rfloor + 1$. Let C be such that $|C| = \frac{nk+j}{2} + 1$ and every coordinate of C is k or 0 , except possibly one. Note that C has at least $\frac{n-j}{2}$ zeros and at most $\frac{n+j}{2}$ non-zeros. Now define

$$\mathcal{F}' := \mathcal{F} \cup \{C\} \setminus \{B\},$$

so that \mathcal{F}' is not a centered family (see Figure 2). We claim that $\text{comp}(\mathcal{F}') < \text{comp}(\mathcal{F})$. We only need to compare the number of subsets of B and C that are contained in \mathcal{F} (or \mathcal{F}'). For a set D and an integer ℓ , write

$$\delta_\ell(D) := \{A \in \mathcal{F} : A \subseteq D, |A| = |D| - \ell\},$$

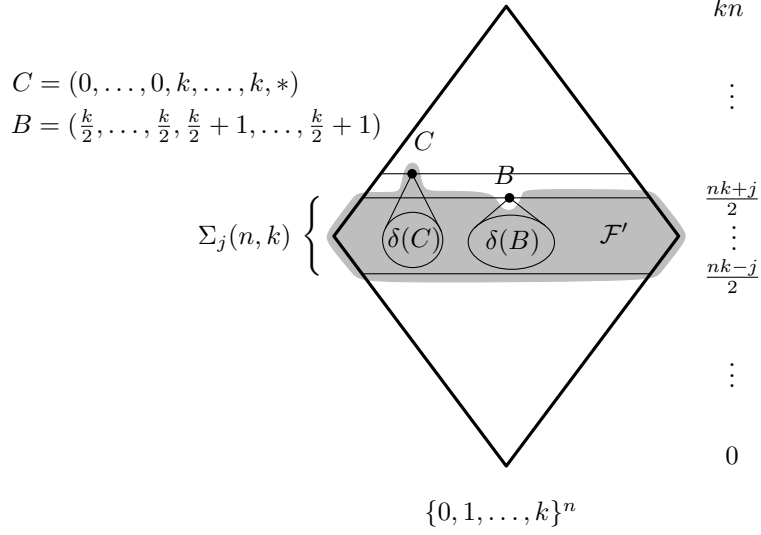


Figure 2: A non-centered family $\mathcal{F}' \subseteq \{0, \dots, k\}^n$ which has smaller number of comparable pairs than the centered family \mathcal{F} .

that is, the collection of subsets of D that are in \mathcal{F} , and are ℓ levels below D . Let

$$\delta(D) = \bigcup_{\ell=0}^n \delta_\ell(D).$$

We have the estimate

$$|\delta(B)| = \sum_{\ell=0}^{j-1} |\delta_\ell(B)| > |\delta_{j-1}(B)| > \binom{n}{j-1}.$$

Note that for $0 \leq \ell \leq j$ we have

$$|\delta_\ell(C)| \leq \binom{\frac{n+j}{2} + \ell - 1}{\ell} = (1 + o(1)) \binom{\frac{n+j}{2}}{\ell},$$

since the right hand side of the first inequality counts the number of non-negative solutions to the equation $a_1 + \dots + a_{(n+j)/2} = \ell$. Hence we get

$$\begin{aligned} |\delta(C)| &= \sum_{\ell=0}^j |\delta_\ell(C)| \leq (1 + o(1)) \sum_{\ell=0}^j \binom{\frac{n+j}{2}}{\ell} \leq 2 \binom{(0.5 + \varepsilon^{3/2})n}{j} \\ &\leq n \cdot (0.5 + \varepsilon^{4/3})^{(1+\varepsilon) \log_2 n} \binom{n}{j-1} < |\delta(B)|, \end{aligned}$$

where the last inequality holds because $(0.5 + \varepsilon^{4/3})^{(1+\varepsilon)} < \frac{1}{2}$ for $\varepsilon < 0.01$. Hence $\text{comp}(\mathcal{F}') < \text{comp}(\mathcal{F})$ and this completes the proof. \square

6 Proof of Theorem 1.7

Recall that $\mathbf{P}(n)$ denotes the collection of posets of order n that are rank-symmetric and rank-unimodal, and let $P \in \mathbf{P}(n)$. Furthermore, recall that $|A|$ denotes the rank of an element $A \in P$, $\text{comp}(A, \mathcal{G}) := |\{B \in \mathcal{G} : B \subset A \text{ or } A \subset B\}|$, and $N_r(A) := \{B : |B| = r, B \subseteq A \text{ or } A \subseteq B\}$.

A poset P of rank n has *property (Q)* if all of the following hold:

- (Q1) If $|B| < |A|$ and $||B| - n/2| < ||A| - n/2|$, then $|N_{|B|+i}(B)| \leq |N_{|A|-i}(A)|$ for every $i \in \{1, \dots, |A| - |B|\}$.
- (Q2) If $|B| > |A|$ and $||B| - n/2| < ||A| - n/2|$, then $|N_{|B|-i}(B)| \leq |N_{|A|+i}(A)|$ for every $i \in \{1, \dots, |B| - |A|\}$.
- (Q3) If $n/2 \leq |B| < |A|$, then $|N_{|B|-i}(B)| \leq |N_{|A|-i}(A)|$ for every $i \geq 1$.
- (Q4) If $n/2 \geq |B| > |A|$, then $|N_{|B|+i}(B)| \leq |N_{|A|+i}(A)|$ for every $i \geq 1$.

The key result of this section is the lemma below, which will easily imply Theorem 1.7.

Lemma 6.1. *If a rank-symmetric and rank-unimodal poset P of rank n has Property (Q), then P has the centeredness property.*

Proof. Let $P \in \mathbf{P}(n)$ that has Property (Q). We say that a family $\mathcal{F} \subseteq P$ is *mid-compressed* if for every comparable pair $(A, B) \in \text{Comp}(\mathcal{F})$ such that $||B| - n/2| < ||A| - n/2|$, $A \in \mathcal{F}$ implies $B \in \mathcal{F}$.

Claim 6.2. *For every $M \in \{1, \dots, |P|\}$, there exists an M -optimal family in P that is mid-compressed.*

Proof. The proof of this claim is essentially the same as Kleitman's proof [7] of Theorem 1.1 and hence similar to our proof of Lemma 2.3, so we only give a sketch here. We show by induction on M that there exists an M -optimal family that is centered. The base case is $M \leq \Sigma_1(n, k)$, in which case there exists an antichain in $\mathcal{L}_{n/2}$ of size M .

Let now $M > \Sigma_1(n, k)$, and define an order relation on the collection of subsets of P of order M by setting $\mathcal{G} < \mathcal{F}$ if

- $\text{comp}(\mathcal{G}) < \text{comp}(\mathcal{F})$, or
- $\text{comp}(\mathcal{G}) = \text{comp}(\mathcal{F})$ and $\sum_{G \in \mathcal{G}} ||G| - n/2| < \sum_{F \in \mathcal{F}} ||F| - n/2|$.

Given a family $\mathcal{F} \subset P$ of size M that is not mid-compressed we will find a family \mathcal{G} of size M that improves \mathcal{F} (that is, $\mathcal{G} < \mathcal{F}$). Since only mid-compressed families cannot be improved this way this will show that there exists an M -optimal mid-compressed family.

Let $\mathcal{F} \subset P$ be a family of size M that is not mid-compressed. Then there exist elements A and B such that $A \in \mathcal{F}$, $B \notin \mathcal{F}$, and $||B| - n/2| < ||A| - n/2|$. W.l.o.g. there exists such a pair with $|A| > n/2$. Among all such pairs (A, B) consider the pairs with $|A|$ is maximal, and then among these pick one with $|B|$ maximal. Note that this implies that whenever $C \in P$ is such that $C \subset A$ and $|C| > |B|$ then $C \in \mathcal{F}$. Moreover whenever $C \in P$ is such that $B \subset C$ and $|C| > |A|$ then $C \notin \mathcal{F}$. Let $a := |A|$ and $b := |B|$.

Form a bipartite graph with vertex sets \mathcal{F}_a and $\overline{\mathcal{F}_b}$ with edges between comparable pairs. If there exists a matching f between \mathcal{F}_a and $\overline{\mathcal{F}_b}$ covering \mathcal{F}_a , then replacing \mathcal{F}_a with the matching elements $f(\mathcal{F}_a)$ does not increase the number comparable pairs in \mathcal{F} (since P has Property (Q1)), but decreases $\sum_{F \in \mathcal{F}} ||F| - n/2|$ and hence improves the family. From now on suppose that there is no such matching. Let $\mathcal{X} = \mathcal{F}_a$ and let \mathcal{Y} be the family of neighbors of \mathcal{F}_a in $\overline{\mathcal{F}_b}$.

Case 1: $|\mathcal{X}| \leq |\mathcal{Y}|$. Since there is no matching between \mathcal{X} and \mathcal{Y} covering \mathcal{X} , we can find a maximal vertex set $\mathcal{X}_0 \subset \mathcal{X}$ such that $|N(\mathcal{X}_0)| < |\mathcal{X}_0|$. Let f be a matching between $\mathcal{F}_a - \mathcal{X}_0$ and $\mathcal{Y} - N(\mathcal{X}_0)$ covering $\mathcal{X} - \mathcal{X}_0$, which exists by the maximality of \mathcal{X}_0 . Then $\mathcal{G} := \mathcal{F} \cup f(\mathcal{X} - \mathcal{X}_0) - (\mathcal{X} - \mathcal{X}_0)$ satisfies $\mathcal{G} < \mathcal{F}$ (again using that P has Property (Q1)).

Case 2: $|\mathcal{X}| > |\mathcal{Y}|$. If there exists a matching f covering \mathcal{Y} then replacing $f(\mathcal{Y})$ by \mathcal{Y} improves \mathcal{F} . Otherwise, let $\mathcal{Y}_0 \subset \mathcal{Y}$ be minimal such that $|N(\mathcal{Y}_0)| < |\mathcal{Y}_0|$. Consider the following two cases:

- a) If there is a matching f between \mathcal{Y}_0 and $N_{\mathcal{X}}(\mathcal{Y}_0)$ covering $N(\mathcal{Y}_0)$, then let $\mathcal{G} := (\mathcal{F} \setminus N_{\mathcal{X}}(\mathcal{Y}_0)) \cup f(N_{\mathcal{X}}(\mathcal{Y}_0))$. Since there is no edge between $f(N_{\mathcal{X}}(\mathcal{Y}_0))$ and \mathcal{F}_a , we have $\text{comp}(\mathcal{G}) < \text{comp}(\mathcal{F})$.
- b) Otherwise, there exists a vertex set $\mathcal{Z} \subseteq N_{\mathcal{X}}(\mathcal{Y}_0)$ with $|N_{\mathcal{Y}_0}(\mathcal{Z})| < |\mathcal{Z}|$. Then $\mathcal{Y}'_0 := \mathcal{Y}_0 \setminus N_{\mathcal{Y}_0}(\mathcal{Z})$ is smaller than \mathcal{Y}_0 and it is easy to check that $|N_{\mathcal{X}}(\mathcal{Y}_0)| < |\mathcal{Y}_0|$, a contradiction with minimality of \mathcal{Y}_0 .

This finishes the proof of the claim that there exists an M -optimal mid-compressed family. \square

From now on we assume that there exists an M -optimal mid-compressed family \mathcal{F}^* that is not centered. Recall that $\Sigma_r(P)$ denotes the total size of the middle r layers of P . Define the integer $j \geq 0$ such that $\Sigma_{j-1}(P) < M \leq \Sigma_j(P)$. Let $\mathcal{G} \subset P$ be the centered family of size $\Sigma_j(P)$ and write $\Delta(\mathcal{G}) := \max\{\text{comp}(A, \mathcal{G}) : A \in \mathcal{G}\}$ for the maximum degree of the graph with vertex set \mathcal{G} and edges corresponding to comparable pairs in P . Let $\text{comp}(M-1) := \min\{\text{comp}(\mathcal{F}) : \mathcal{F} \subseteq P, |\mathcal{F}| = M-1\}$. The following statement is very similar to Claim 2.8:

Claim 6.3. *We have $\text{comp}(\mathcal{F}^*) \leq \text{comp}(M-1) + \Delta(\mathcal{G})$.*

Proof. It suffices to construct a family \mathcal{F} of size M with at most $\text{comp}(M-1) + \Delta(\mathcal{G})$ comparable pairs. As \mathcal{F}^* is M -optimal it contains at most this many comparable pairs. By induction we know there exists a centered $(M-1)$ -optimal family \mathcal{H} . Since $\mathcal{H} \subset \mathcal{G}$, adding to it any element of $\mathcal{G} \setminus \mathcal{H}$ increases the number of comparable pairs by at most $\Delta(\mathcal{G})$. \square

Since \mathcal{F}^* is not centered, it contains an element A such that for all elements $B \in \mathcal{G}$ we have $||A| - n/2| > ||B| - n/2|$. Since \mathcal{F}^* is mid-compressed and P has properties (Q3) and (Q4), this implies that $\text{comp}(A, \mathcal{F}^*) \geq \Delta(\mathcal{G})$. Hence $\text{comp}(\mathcal{F}^*) \geq \text{comp}(M-1) + \Delta(\mathcal{G})$. By Claim 6.3 this implies that every family of size M contains at least $\text{comp}(M-1) + \Delta(\mathcal{G})$ comparable pairs. As shown in the proof of Claim 6.3 this value can be achieved by a centered family, completing the proof of Lemma 6.1. \square

One well-known poset that satisfies the assumptions of Lemma 6.1 is the Boolean lattice $\mathcal{P}(n)$. Therefore, Lemma 6.1 implies Theorem 1.1 – rather unsurprisingly since the proof of Lemma 6.1 was motivated by Kleitman’s proof of Theorem 1.1.

Let q be a prime power and let $n \geq 1$. To finish the proof of Theorem 1.7, we only need to check that the assumptions of Lemma 6.1 hold for $\mathcal{V}(q, n)$.

Claim 6.4. *$\mathcal{V}(q, n)$ is rank-symmetric.*

Proof. The map $V \rightarrow \mathbb{F}_q^n \setminus V$ takes the set of subspaces of dimension k into the set of subspaces of dimension $n-k$ bijectively. \square

Claim 6.5. *$\mathcal{V}(q, n)$ is rank-unimodal.*

Proof. Note that the number of subspaces of $\mathcal{V}(q, n)$ of dimension k , written as $\begin{bmatrix} n \\ k \end{bmatrix}_q$, can be expressed as (see e.g. [10]):

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{[n]!}{[k]![n-k]!},$$

where

$$[n]! = [1] \cdot [2] \cdot \dots \cdot [n], \quad \text{and} \quad [i] = q^i - 1.$$

Rank-unimodality of $\mathcal{V}(q, n)$ is easily seen to follow from this formula. \square

Claim 6.6. $\mathcal{V}(q, n)$ has Property (Q).

Proof. Properties (Q1)–(Q4) follow from the observation that if S is a subspace of \mathbb{F}_q^n of dimension m then the number of spaces $S' \subset S$ of dimension $m - k$ is $\begin{bmatrix} m \\ k \end{bmatrix}_q$ and the number of spaces S' with $S \subset S'$ and $\dim(S') = m + k$ is $\begin{bmatrix} n-m \\ k \end{bmatrix}_q$. \square

7 Open problems

Recall that \mathbf{P} is the collection of posets that are rank-symmetric and rank-unimodal and let $\mathbf{C} \subset \mathbf{P}$ be the collection of posets which have the centeredness property. The main open problem that this paper has only barely begun to explore asks for an easy way to decide whether a poset $P \in \mathbf{P}$ is in \mathbf{C} . We know that $\{0, 1\}^n \in \mathbf{C}$ and $\mathcal{V}(q, n) \in \mathbf{C}$ but for $k \geq 2$ and n large we have $\{0, 1, \dots, k\}^n \in \mathbf{P} \setminus \mathbf{C}$.

Now let P_G be the lattice of subgroups of a finite Abelian group G . It was shown in [4] that P_G is rank-unimodal. The following general question is likely to be difficult to solve in full generality but any progress would be interesting.

Question 7.1. *For what Abelian groups G is it true that $P_G \in \mathbf{C}$?*

Observe that most results of this paper are special cases of Question 7.1:

- if $G = C_{p_1} \times C_{p_2} \times \dots \times C_{p_n}$ for distinct primes p_1, p_2, \dots, p_n then P_G is (isomorphic to) the Boolean lattice and hence $P_G \in \mathbf{C}$,
- if $G = C_{p_1^k} \times C_{p_2^k} \times \dots \times C_{p_n^k}$ for distinct primes p_1, p_2, \dots, p_n then P_G is isomorphic to the lattice $\{0, 1, \dots, k\}^n$ under inclusion and hence if $n \geq n_0(k)$ then $P_G \in \mathbf{P} \setminus \mathbf{C}$.
- if $G = (C_p)^n$ for p prime then P_G is isomorphic to $\mathcal{V}(p, n)$ and hence $P_G \in \mathbf{C}$.

Question 7.1 can be asked for other members of \mathbf{P} , see e.g. [10]. A natural generalization of the centeredness property is as follows. For an integer $r \geq 2$ say that a poset $P \in \mathbf{P}$ has the *r-centeredness property* if for all M with $0 \leq M \leq |P|$, among all families $\mathcal{F} \subset P$ of size M , the number of r -chains contained in P is minimized by a centered family. Denote the collection of posets with the r -centeredness property by \mathbf{C}_r and note that $\mathbf{C} = \mathbf{C}_2$. A long-standing conjecture in this area due to Kleitman [7] is that $\{0, 1\}^n \in \mathbf{C}_r$ for all n, r . For recent progress on this conjecture we refer the reader to [3, 5, 6]. Asking for a characterisation of \mathbf{C}_r is currently out of reach, but finding interesting necessary and/or sufficient conditions for a poset $P \in \mathbf{P}$ to be in \mathbf{C}_r could be a fine result.

In a different direction one could improve Theorem 1.3 and investigate further for which M Conjecture 1.2 holds.

Question 7.2. *For which k and M does there exist an M -optimal centered family in $\{0, 1, \dots, k\}^n$?*

The same question can be asked for ‘centered’ replaced by ‘canonical centered’ (i.e. centered families with at most one partially filled layer). We expect that for $k = 2$ the answer to Question 7.2 contains the interval $[0, \Sigma_5(n, 2)]$. It seems plausible that for $M \leq \Sigma_{\log_2 n}(n, k)$ the centered families are not too far from being best possible, but for much larger M we do not even have a guess what the best families could be. The following question is open whenever \sqrt{n} is replaced by any value between $\log_2 n$ and n .

Question 7.3. *Let $M = \Sigma_{\sqrt{n}}(n, 2)$. What do the M -optimal families in $\{0, 1, 2\}^n$ look like?*

References

- [1] N. Alon, S. Das, R. Glebov, B. Sudakov. Comparable pairs in families of sets. *Journal of Combinatorial Theory, Series B*, 115, 164–185, 2015.
- [2] I. Anderson. Combinatorics of finite sets, *Courier Corporation*, 1987.
- [3] J. Balogh, A. Z. Wagner. Kleitman’s conjecture about families of given size minimizing the number of k -chains, <https://arxiv.org/abs/1609.02262>.
- [4] L. M. Butler. A unimodality result in the enumeration of subgroups of a finite abelian group, *Proc. Amer. Math. Soc.*, 101(4), 771–775, 1987.
- [5] S. Das, W. Gan, B. Sudakov. Sperner’s Theorem and a Problem of Erdős, Katona and Kleitman, *Combin. Probab. Comput.*, 24(4), 585–608, 2015.
- [6] A. P. Dove, J. R. Griggs, R. J. Kang, J-S. Sereni. Supersaturation in the Boolean lattice, *Integers*, 14A, 1–7, 2014.
- [7] D. Kleitman. A conjecture of Erdős–Katona on commensurable pairs of subsets of an n -set, *Theory of Graphs*, Proc. Colloq. held at Tihany, 1966, Akadémiai Kiadó, Budapest 215–218, 1968.
- [8] D. Kleitman. Collections of subsets containing no two sets and their union, *Proceedings of the LA Meeting AMS (1968)*, 153–155, Pub. AMS, 1971.
- [9] J. Noel, A. Scott, B. Sudakov. Supersaturation in Posets and Applications Involving the Container Method, <https://arxiv.org/abs/1610.01521>.
- [10] R. P. Stanley. Log-concave and unimodal sequences in algebra, combinatorics, and geometry, *Ann. New York Acad. Sci.*, 576(1), 500–535, 1989.